IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 10, 2021, accepted December 2, 2021, date of publication January 4, 2022, date of current version February 7, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3140287

Hybrid Deep Learning Model for Fault Detection and Classification of Grid-Connected Photovoltaic System

MOATH ALRIFAEY, (Graduate Student Member, IEEE), WEI HONG LIM, CHUN KIT ANG, ELANGO NATARAJAN, MAHMUD IWAN SOLIHIN, MOHD RIZON MOHAMED JUHARI, AND SEW SUN TIANG

Faculty of Engineering, Technology and Built Environment, UCSI University, Kuala Lumpur 56000, Malaysia

Corresponding authors: Wei Hong Lim (limwh@ucsiuniversity.edu.my) and Chun Kit Ang (angck@ucsiuniversity.edu.my)

This work was supported by the Ministry of Higher Education Malaysia through the Fundamental Research Grant Scheme under Project Proj-FRGS/1/2020/TK0/UCSI/02/4.

ABSTRACT Effective fault detection and classification play essential roles in reducing the hazards such as electric shocks and fire in photovoltaic (PV) systems. However, the issues of interest in fault detection and classification for PV systems remain an open-ended challenge due to manual and time-consuming processes that require the relevant domain knowledge and experience of fault diagnoses. This paper proposes a hybrid deep-learning (DL) model-based combined architectures as the novel DL approaches to achieve the real-time automatic fault detection and classification of a PV system. This research employed the wavelet packet transform (WPT) as a data preprocessing technique to handle the PV voltage signal collected and feeding them as the inputs for combined DL architectures that consist of the equilibrium optimizer algorithm (EOA) and long short-term memory (LSTM-SAE) approaches. The combined DL architectures are able to extract the fault features automatically from the preprocessed data without requiring any previous knowledge, therefore can override the traditional shortages of manual feature extraction and manual selection of optimal features from the extracted fault features. These desirable features are anticipated to speed up the fault detection and classification capability of the proposed DL model with higher accuracy. In order to determine the performance of the proposed fault model, we carried out a comprehensive evaluation study on a 250-kW grid-connected PV system. In this paper, symmetrical and asymmetrical faults have been studied involving all the phases and ground faults such as single phase to ground, phases to phase, phase to phase to ground, and three-phase to ground. The simulation results validate the efficacy of the proposed model in terms of computation time, accuracy of fault detection, and noise robustness. Comprehensive comparisons between the simulation results and previous studies demonstrate the multidisciplinary applications of the present study.

INDEX TERMS Deep distributed energy, equilibrium optimizer algorithm (EOA), fault detection and classification, grid-connected photovoltaic systems, optimal feature selection, wavelet packet transform (WPT).

NOMENCLATURE TABLE

SYMBOL	DEFINITION
ABC-G	Fault of three phase (A,B, and C) to ground (G)
AI	Artificial intelligence
ANFIS	Adaptive neuro-fuzzy inference system
CBM	Condition-based maintenance
CM	Corrective maintenance
CNN	Convolutional neural network
CAE	Contractive autoencoder
db4	Daubechy mother wavelet
dB	Decibel
DEOA	Deep equilibrium optimizer algorithm
DG	Distribution generation
DL	Deep learning

The associate editor coordinating the review of this manuscript and approving it for publication was Syed Islam.

DWT	Discrete wavelet transform
ELM	Extreme learning machine
EOA	Equilibrium optimizer algorithm
GCPV	Grid-connected photovoltaic
GHz	Gigahertz
GP	Generation probability
GPU	Graphics processing unit
GW	Gigawatt
Hz	Hertz
IEEE	Institute of electrical and electronics engineers
IRENA	International Renewable Energy Agency
kHz	Kilohertz
km	Kilometer
kV	Kilovolt
kVA	Kilovolt-ampere
kW	Kilowatt
LSTM	Long short-term memory
MLP	Multilayer perceptron
MOPSO	Multiobjective particle swarm optimization
MVA	Megavolt-ampere
ms	Millisecond
MW	Megawatts
PCC	Point of common coupling
PSM	Preventive scheduled maintenance
PV	Photovoltaic
RPNN	Ridgelet probabilistic neural network
s	Seconds
SAE	Stacked autoencoders
SCADA	Supervisory control and data acquisition
SLT	Slantlet transform
SNR	Signalto noise ratio
ST	Stockwell transform
SV	Support vector
SVM	Support vector machine
T1	Transformer number one
WP	Wave packet
WPT	Wavelet packet transform

I. INTRODUCTION

The importance of renewable energy sources has gained more attention in recent years due to rapidly growing concerns on environmental pollution issues [1], [2]. Despite the economic slowdown brought by the COVID-19 pandemic, the additional global renewable energy capacity generated in the year 2020 has exceeded the earlier estimation and all historical records. Based on the latest data issued by International Renewable Energy Agency (IRENA), the world has increased more than 260 GW (gigawatt) of renewable energy capacity in 2020, i.e., around 50% higher than the energy generation in 2019 [3].

Among all existing types of renewable energy sources, the utilization rates of both solar and wind energy sources have increased dramatically. The power generation of solar-based has gained significant attention worldwide because of the environmental and economic benefits offered. At the end of 2019, a cumulative volume of solar power of 629 GW was settled around the world. The top country for solar power in early 2020 was China, with 208 GW [4], [5], which accounts for one-third of worldwide installed solar capabilities. By the end of 2020, not less than 37 countries around the world have generated a cumulative photovoltaic (PV) capacity of more than 1 GW. Despite the significant environmental and economic benefits offered, the existing PV systems are subjected to numerous faulty issues due to the harsh outdoor operating environment. The undesirable issues such as power losses and potential safety hazards might occur if these faulty issues remain unnoticed and hidden in the PV system [3]. It is of the utmost importance to design effective protection techniques that can monitor and diagnose the occurrence of faults in a PV system. Substantial efforts have been attempted in literature to provide safe and secure protection schemes for the grid-connected PV system.

Recently, different artificial intelligence (AI) techniques were incorporated as the core methodologies of PV fault detection and classification due to their excellent capabilities in addressing feature extraction and classification problems. Some notable AI approaches such as the convolutional neural network [6] and sequential probabilistic neural network [7] were applied to improve the accuracy of fault classification in a PV system. Apart from fault classification, some supervised machine learning techniques were also applied to focus on the feature extraction process in [8], whereas an adaptive neuro-fuzzy inference system (ANFIS) was used in [9] to address the tracking and detection of faulty issues in PV. For this reason, many computational fault detection models based on machine learning algorithms were proposed [7], [8], [10]–[12]. Substantial amounts of literature studies revealed that the availability of labeled fault data is one of the key factors that enable these intelligence-based algorithms to perform effective fault diagnosis for PV systems [12]. For instance, a combination of slantlet transform (SLT) and ridgelet probabilistic neural network (RPNN) was offered to detect and classify the grid faults in a grid-connected PV system [13].Given the properties of SLT and the decomposed waveform features together with the RPNN algorithm, important information or features can be extracted from a fault signal and used to determine the type of fault that occurred in the islanding problem. Numerical features, for instance, the energy, mean, minimum and maximum values, standard deviation, and log energy entropy for the decomposition coefficients of the distorted signal, were determined and used as input dataset features for RPNN. Support vector machine (SVM) has gained its increasing popularity in the machine learning domains because of the commendable and superior performances demonstrated against to other machine learning models. A hybrid framework of SVM with a wavelet multi-resolution singular spectrum entropy was considered for the fault detection and classification in a Canadian system [14]**.** Notably, the training process of SVM model was formulated as a complex quadratic programming

problem that could incur high computational complexity and computing time cost. This demerit could restrict the practicability of SVM model in addressing the real-world fault detection and classification issues that involved the presence of large datasets.

Fault location approaches share an essential role in the management of outages and faults, subsequently the safety of electrical supply [15]. In [16] and [17], different machine learning models such as multilayer perceptron (MLP) and extreme learning machine (ELM) were combined with a signal processing technique known as discrete wavelet transform (DWT) to improve the performances in detecting, classifying, and locating the fault events in the radial distribution grid. In [18], Aljohani *et al.* proposed a hybrid technique based on the machine learning and Stockwell transform (ST) to detect, locate and classify the single-line-to-ground fault in a modeled distribution feeder. In [19], an improved ST technique was used with machine learning to extract useful features to locate and identify the fault section. In [15], an impedance-based equation with recorded voltage and current was designed as a new fault location approach to identify the fault sections of power distribution networks incorporated with distributed generations. However, this approach is not applicable in various conditions and modes. In [20], a novel method of impedance-based was proposed to identify the precise fault location, in terms of fault point and distance, in the distribution networks of electric energy. One of the main drawbacks of this methodology is the necessity of categorizing the faults before being able to define their locations. This is for the reason that the resources of distributed generation are currently broadly implemented in the distribution networks and might not be neglected.

Over the years, different combined architectures of deep learning approaches were covered extensively in the literature. The combined DL architectures have the ability of automatic learning and accurately define the deep features of faults that are generally unlabeled and cannot be precisely identified using the traditional fault approaches. Several approaches are reported in the literature to address this issue. For example, research has provided evidence for a combination of two deep learning architectures (i.e., SAE and LSTM) to detect the condition of an anomaly in a fully unsupervised mode [21]. Li *et al.* [21] also combined the SAE-LSTM with wavelet packet decomposition (WPD) to examine and validate the anomaly detection performance in a rotating machine. In [22], Tovar *et al.* proposed a hybrid model of CNN-LSTM (convolutional neural network) for predictions of PV power, where the CNN method was used to extract and select the local features, while LSTM was used for extracting the temporal features of the real data in Mexico Temixco. The hybrid model of CNN-LSTM introduced by [22] has better prediction than the single prediction model. In [23], Ahmadipour *et al.* built a novel technique by combining the wavelet packet transform (WPT) and probabilistic neural network (PNN) for islanding detection in a PV system, where WPT was applied to extract the islanding feature and PNN was used for islanding detection. In [24], Kumar *et al.* studied the detection and diagnosis of online PV faults based on Wavelet Packets and the instantaneous measurement to distinguish the partial shading condition from faults for overcoming false fault detecting and system falling. However, this method has the disadvantage of not being applicable for large plants, especially when the irradiance amongst PV modules is dissimilar.

Apart from the above approaches, the idea of using SAE to extract features in deep learning architecture has also been deliberated in other studies. In [25], Gao *et al.* built the DL architecture based on the stacking of SAE and an improved multi-grained cascade forest that can perform the automatic feature extraction for the diagnosis of multiple faults. In [26], Chen *et al.* built a model that combined the SAE and contractive autoencoder (CAE) to enhance the ability of data learning and the strength of feature extraction of autoencoder in DL architecture. Chen *et al*. also employed the quantum ant colony algorithm to contribute to the parameter optimization of the particular location and used multiple autoencoders to achieve better performance in planetary gear fault diagnosis. In [27], Qi *et al.* proposed an enhanced diagnosis technique based on SAE, and the results show that the proposed SAE is able to extract extra distinctive high-level features and deliver superior performance in fault diagnosis when compared with the conventional ML methods. Referring to be abovementioned works, it can be observed that the combinations of SAE, LSTM, and WPT methods in DL architecture have demonstrated promising performance to improve the accuracy of fault detection and diagnosis in application domains.

Other protection techniques were also proposed to solve the fault detection and classification issues of grid-connected PV systems based on Thevenin Equivalent Resistance [28], radial basis function networks [29], wavelet transform [24], statistical monitoring [30], fuzzy inference [31] and etc. Furthermore, several studies were proposed to diagnose the PV faults used in different applications by first identifying the root causes of these faulty situations and then proposing the corresponding solutions to prevent their occurrence in the future, including the fault diagnosis [32], [33], prediction [34], cost irradiance monitoring [35], maintenance [10], and islanding [36]–[38]. Fault monitoring and diagnosing are the effective maintenance mechanisms used by most existing PV systems to prevent the occurrence of faulty issues and corresponding consequences. The initial detection and classification of failure modes are also useful for maintenance applications in order to identify the hazard patterns for maintenance reliability and machine availability [39], [40].

Despite having the substantial amounts of related works published in this research area, there is still lacking the robust methods that are able to detect and classify the PV faults accurately with low computational time [41], [42].

Nevertheless, most of these existing methods encountered the challenges of attaining high accuracy of fault detection and classification because of the manual feature extraction process that is not only time-consuming but also

requires substantial amounts of domain knowledge in the PV fault diagnosis. Furthermore, most of these existing methods are also observed not able to overcome the challenges of manual feature extraction and feature selection simultaneously [43]–[45].

The research gaps observed from previous literature, such as the limitation of manual feature extraction and selection as well as the time-consuming process of PV fault classification, have motivated the current study to develop a new hybrid deep learning (DL) model that aims to achieve better accuracy of PV fault classification within shorter computational time. In particular, several intelligence methods known as the wavelet packet transform (WPT), the deep-stacked encoder (SAE), equilibrium optimizer algorithm (EOA), and long short-term memory (LSTM) are hybridized into the proposed DL model, aiming to obtain an optimal feature subset required for the unsupervised fault classification in achieving the maintenance scheduling and failure prevention of PV system.

In contrary to many existing works, the proposed DL model is unique because it aims to fill the gaps for all three stages (i.e., features extraction, optimal feature selection, and fault classification) simultaneously when learning and optimizing the fault detection model for PV electrical units even under the presence of the noisy environment. These competitive advantages enable the proposed hybrid DL model to perform the fault detection and classification of PV plants in faster and more reliable manners without leaving them unnoticed, which in turn can result in undesirable consequences such as power losses and safety hazards in the PV systems.

To that end, the most significant contributions of the present study are summarized as follows:

- A novel hybrid DL model is proposed to tackle the PV fault detection problems, where the key challenges of feature extraction, optimal feature selection, and fault classification encountered on the proposed hybrid DL model are addressed simultaneously via strategic hybridization of WPT, SAE, LSTM, and EOA methods.
- Hybrid mathematical methods are proposed to extract and select the optimal features required for PV fault detection in addressing the drawbacks of manual feature selection and excessive time incurred in the training process of PV fault classification and detection.
- The proposed hybrid DL model enables unsupervised learning from the unlabeled PV measured signals, implying that the proposed approach is not only able to extract the significant fault features automatically but also can optimize the number of selected features and processing time of fault classification.
- The proposed hybrid DL model is expected to contribute to PV researches, particularly in the aspects of feature extraction, feature selection, and fault classification. It is envisioned that the same proposed method can be applied in fault diagnosing and prevention for other industrial applications, even in a noisy environment.
- An extensive study is conducted to evaluate the effects of high-level noises on the protection performance offered by the proposed hybrid DL model. This analysis is crucial to verify the applicability of the proposed method to be deployed in the actual microgrid system, i.e., an issue that was frequently overlooked in the previous studies.

The rest of the paper is structured as follows: Section II presents the proposed methodology of the hybrid DL model. Section III describes the data processing and mechanisms incorporated into each subsystem of the proposed hybrid DL model, followed by Section IV that outlines the case study used to evaluate the performance of the proposed hybrid DL method. Section V presents the simulation results and discussion used to validate the proposed hybrid DL model. Finally, conclusions are drawn in Section VI.

II. THE OVERALL ARCHITECTURE OF THE PROPOSED METHODOLOGY

In this section, a hybrid DL model is proposed to achieve better performances in terms of fast PV fault detection with feature reduction capability by combining the wavelet packet transform (WPT), deep-stacked autoencoders (SAE), and equilibrium optimizer algorithm and long short-term memory (EOA-LSTM) approaches. Fig.1 displays the proposed hybrid DL model used for fast PV fault detection and classification with reduced features.

Data preprocessing is first performed on the initial data set because this step is crucial to ensure the effectiveness and accurateness of data mining models [46]. It is suggested in the literature that different methods can be used to initialize and preprocess the dataset in order to attain better classification performance.

Firstly, the WPT method is applied for data preprocessing of the collected PV voltage signal. Secondly, an SAE is employed to extract the fault features automatically from the preprocessed data, aiming to override the shortages of traditional manual feature extraction. Then, the EOA algorithm is applied to choose the optimal features automatically from the extracted fault features without requiring any previous knowledge. An EOA algorithm is constructed with multiple LSTM to speed up the fault detection and classification as well as to enhance the feature learning capability of the LSTM model. Finally, the LTSM is adopted as a PV fault classification model, in which the dataset of extracted features is divided into two portions. Specifically, 80% of the dataset is used for training the intelligent LTSM model, and the remaining 20% of the dataset is used for testing the performance of LTSM in fault detection and classification. After that, the trained model can be deployed to alert decision-makers for preparing the appropriate maintenance tasks in order to prevent hazardous faults in the future.

III. DATA AND METHODS

A. DATA PREPROCESSING

Wavelet transform is commonly used as a signal processing tool for many power system applications due to its ability to

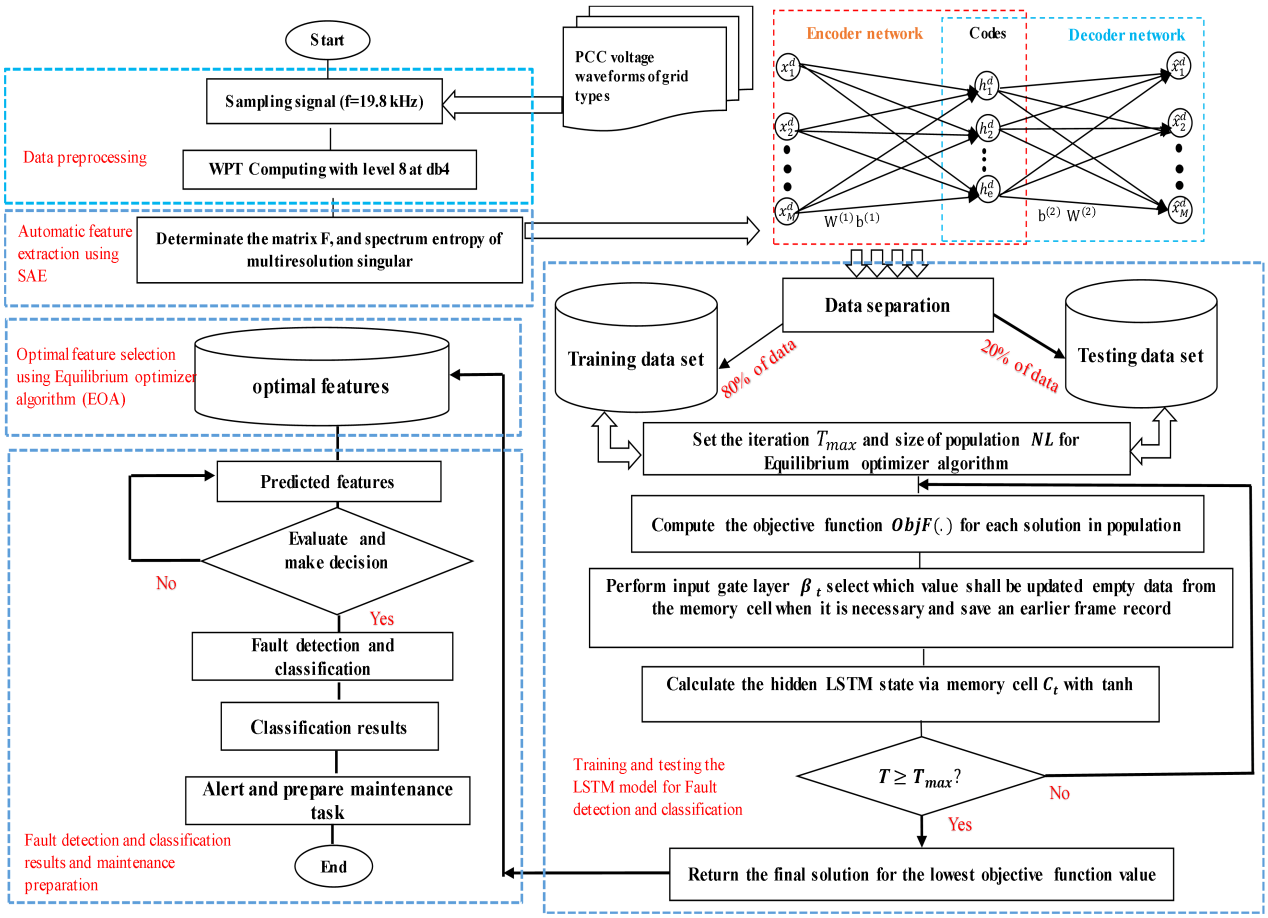

FIGURE 1. Proposed hybrid DL model for PV fault detection and classification.

detect the non-stationary transients encountered during signal measurement. Nevertheless, wavelet transform tends to suffer from some drawbacks such as the low immunity for noise and disturbances, batch processing, and varying frequency sub-band. In order to address the challenges brought by wavelet transform, wavelet packet transform (WPT) is used as a signal processing tool in this paper because it can offer more robust noise immunity with signal properties for enhanced feature extraction and time resolution [47].

In contrary to wavelet transform, WPT has demonstrated its promising capability to analyze the local discontinuities within an input signal that might contain more significant features for fault diagnosis. For this reason, WPT is considered as a more appropriate signal processing tool for handling the non-stationary and steady-state during the data ingestion process. In general, the decomposition depends on the sampling frequency; for this study, the sampling rate is set as 19.8 kHz (Kilohertz) because we were simulating the system according to UK standards and the study by [14]. The 8-level is selected to cover all frequency bands in order to have a more evident observation on the band that has arisen fault. According to the experimental findings of [13], [48], daubechy mother wavelet (db4) is proven suitable to be employed as the mother wavelet because of its high performance, i.e., high precision detection with short processing time. Other relevant fault parameters are selected according to the operative requirements in IEEE 1547 standard, the practicable topology of the operational network, and testing practices are recommended by protection relay manufacturers. More details of technical operations for the electrical system considered in the current case study can be found in [23].

During the decomposition process, the original input signal can be divided into the low-frequency coefficient of A1 and the high-frequency coefficient of D1. It is noteworthy that the conventional orthogonal wavelet method only considers low-frequency regions during the decomposition process and neglects the high-frequency regions that might contain the significant features that appear in these local discontinuities of the signal. On the contrary, the WPT structure can overcome the deficiency of the orthogonal wavelet method by providing a well-adjusted structure analysis and precise frequency resolution. In this paper, WPT is used to provide the time-frequency function of PV signals, enabling the more effective detection of faulty events. Equations (1) to (8) are presented to provide further explanations on mathematical formulations involved in WPT. The computation of wave

packet functions(WP) $W^i_{l,k}$ can be expressed as follows [49]:

$$W^i_{l,k}(t) = \sqrt{2^l} W^i(2^l t - k) \tag{1}$$

where l and k are the scales and translation parameters respectively, and i is the oscillation index $\left(i = 0, 1, \ldots 2^l\right)$.

The first two WP functions for $i = 0$ and $i = 1$ are shown in (2) and (3), respectively, i.e.,

$$\varphi(t) = W^0_{0,0}(t) = W^0(t) \tag{2}$$

$$\psi(t) = W^1_{0,0}(t) = W^1(t) \tag{3}$$

where $\varphi(t)$ and $\psi(t)$ denote the scaling function and mother wavelet, respectively, where $l = k = 0$. Meanwhile, the remaining WP functions with $i = 2, 3, \ldots, 2^l$ (where 2^l is the number of decomposed packets) are expressed as follows:

$$W^{2i}(t) = \sum_k h(k) W^i_{1,k}(t) \tag{4}$$

$$W^{2i+1}(t) = \sum_k g(k) W^i_{1,k}(t) \tag{5}$$

where both $h(k)$ and $g(k)$ represent the functions of low-pass and high-pass filters, respectively, that are defined as follows:

$$h(k) = 1/\sqrt{2}\langle\varphi(t), \varphi(2t - k)\rangle \tag{6}$$

$$g(k) = 1/\sqrt{2}\langle\psi(t), \psi(2t - k)\rangle \tag{7}$$

The above-mentioned low-pass and high-pass filters are orthogonal, where $g(k) = (-1)^k h(1 - k)$. According to (6) and (7), the following WP functions can therefore be further expressed as:

$$W^{2i}(t) = \sqrt{2}\sum_k h(k) W^i(2t - k) \tag{8}$$

$$W^{2i+1}(t) = \sqrt{2}\sum_k g(k) W^i(2t - k) \tag{9}$$

Generally, the wavelet packet coefficient of $C^i_l(k)$ for a continuous signal of $x(t)$ is represented as:

$$C^i_l(k) = \langle x, W^i_{l,k}\rangle = \int_{-\infty}^{+\infty} x(t) W^i_{l,k}(t)\, dt \tag{10}$$

For discrete signal, the coefficients of WPT decomposition can be calculated as:

$$C^{2i}_{l+1}(\tau) = \langle x, W^i_{l,k}\rangle = \sum_k h(k - 2\tau) C^i_l(k) \tag{11}$$

$$C^{2i+1}_{l+1}(\tau) = \langle x, W^i_{l,k}\rangle = \sum_k g(k - 2\tau) C^i_l(k) \tag{12}$$

Classical WPT-based measures can produce various features such as entropy, energy, standard variation, etc., from the WP coefficients obtained using (10) to (12). Energy and entropy are the most popular measurements that can be applied to determine the irregular features of normal and faulty signals. Both the energy and entropy can vary with the frequency and time components of different signals and conditions. In this paper, the energy characteristics of WPT nodes are considered to perform fault detection and classifications. Both of the energy and entropy measures can be determined through the coefficients of WPT extracted from (10)-(12). The Shannon entropy is calculated as:

$$Energy_{l,i} = -\sum_{n=1}^{N} \left|C^i_{l,k}(n)\right|^2 log \left|C^i_{l,k}(n)\right|^2 \tag{13}$$

The logarithmic energy entropy can be determined as:

$$Energy_{l,i} = \sum_{n=1}^{N} log\left[\left(C^i_{l,k}(n)\right)\right]^2 \tag{14}$$

where N denotes the sampling point number; $C^i_{l,k}$ is the coefficients extracted from the l-th level of the i-th frequency band with $i = 2^l - 1$.

B. FAULT FEATURE EXTRACTION

It is essential to design an effective feature extraction process from the measurement signals so that the proposed method is able to accurately discriminate different kinds of faults that appear in the PV system. In this paper, an unsupervised deep learning model known as stacked autoencoder (SAE) is employed as the feature extraction method because of its promising performances in extracting important information from the signals of condition monitoring system concerning the analysis difficulties [50]. SAE has the excellent capability of performing dimensionality reduction on the input signal into any desired sizes by leveraging its hidden layer as a feature extractor, as well as can predict the output and the same input data without requiring labels. For this reason, SAE is able to automatically extract the significant fault features from input signals without requiring any data labeling and yet can offer better descriptions of fault features than the original data [51]. Fig. 2 shows the network structure of a three-layer SAE (such as input layer, hidden layer, output layer). The mapping and reconstructing processes of SAE can be attained through the encoder (i.e., from input vector into a hidden vector) and decoder (i.e., from hidden vector into output vector) networks, respectively.

Suppose that a training sample set denoted as $x^d = \left[x^d_1, x^d_2, \cdots, x^d_M\right]^M \in R^M$ is mapped into the hidden layer features by (15), where $d = 1, 2, \ldots, M$ and M represent the number of samples in the dataset $\left\{x^d\right\}^M_{d=1}$. Denote $h^d = [h^d_1, h^d_2, \cdots, h^d_e]^T \in R^e$ as the feature vector of the hidden map, where e represents a dimension of hidden unit, it is then calculated as:

$$h^d = f(W^{(1)}x^d + b^{(1)}) \tag{15}$$

where $f(\cdot)$ is an activation function; $W^{(1)}$ and $b^{(1)}$ are two matrices containing the weight and bias values between the input and hidden layers, respectively. During the decoding process, the feature vector of the hidden layer is transformed into an output vector of $\hat{x}^d = \left[\hat{x}^d_1, \hat{x}^d_2, \cdots, \hat{x}^d_M\right]^M \in R^M$, i.e.,

$$\hat{x}^d = f(\mathrm{W}^{(2)}h^d + b^{(2)}) \tag{16}$$

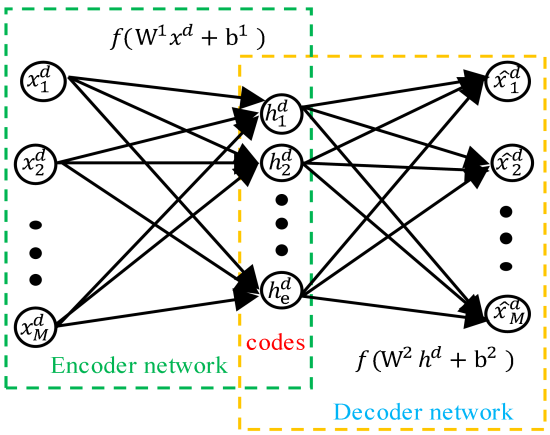

FIGURE 2. SAE neural network Structure.

where $W^{(2)}$ and $b^{(2)}$are two matrices containing the weight and bias values between the hidden and output layers, respectively. The reconstruction error $L(x^d, \hat{x}^d)$ between $\hat{x}^d$ and x^d is then expressed as follow:

$$L\left(x^d, \hat{x}^d\right) = \frac{1}{2}\left\|x^d - \hat{x}^d\right\|^2 \tag{17}$$

The overall cost function of SAE is depicted as:

$$J(W, b) = \left[\frac{1}{M}\sum_{d=1}^{M} L\left(x^d, \hat{x}^d\right)\right] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}\left(W_{ji}^{(l)}\right)^2 \tag{18}$$

where λ denotes the decay weight parameter; $W_{ji}^{(l)}$ is a linking weight for neurons i into j with layer $(l, l+1)$; n_l and s_l are the network layer number and neuron number for layer l, respectively. Notably, the second term in (18) is used to measure the differences between the input vector of all training samples and their corresponding reconstructed vectors.

Suppose that $\hat{\rho}_g$ is the average value for activating the g-th hidden unit, where g denotes the neuron index in the hidden layer with $g = 1, 2, \cdots, e$. The averaging for activated values is calculated as follow:

$$\hat{\rho}_{\mathrm{g}} = \frac{1}{M}\sum_{\mathrm{g}=1}^{e} h_{\mathrm{g}}^d \tag{19}$$

Furthermore, the network weight and bias values are adjusted to avoid over-fitting during the training process using gradient descent. Therefore, the cost function of SAE can be redefined as follow:

$$J_{sparse}(W, b) = J(W, b) + \theta\sum_{g=1}^{e}\left(\rho\log\frac{\rho}{\hat{\rho}_g} + (1-\rho)\log\frac{1-\rho}{1-\hat{\rho}_g}\right) \tag{20}$$

where θ is the sparsity penalty parameter, which is used to control the relative significances for reconstruction and penalty; ρ denotes a sparsity parameter.

C. OPTIMAL FEATURE SELECTION

A deep equilibrium optimizer algorithm (DEOA) is proposed to select the optimal combinations of features used by the proposed hybrid DL model to perform real-time automatic detection and classification of PV faults. Equilibrium optimization (EOA) was developed in [52] based on the equilibrium of dynamic mass for control volume by seeking the balanced states of the system in order to solve optimization problems. EOA is a competitive optimization algorithm due to its several advantages, such as the good balancing of exploration and exploitation searches, simplicity of implementation, and good population diversity. Recently, EOA has been successfully applied across different fields such as network reconfiguration of the power system [53], image segmentation [54], dynamic model of the fuel cell [55], parameter estimation of PV cell [56], [57], and automatic voltage regulator system [58]. Despite the benefits offered, EO has the drawbacks of lacking attention on fitness assignment, and it is not able to satisfy the contradictory goals brought by the multiobjective functions simultaneously due to its high tendency of reaching equilibrium in one objective and failing in the remaining objectives.

In order to overcome the drawbacks of EOA, the DEOA is proposed in this paper to deal with optimal feature selection that is commonly formulated as a multiobjective optimization problem. The proposed DEOA employs a hyper learning approach that can leverage the concepts of personal best and personal worst states during the solution updating process in solving the multiobjective feature selection problem. Without loss of generality, DEOA consists of n sub-swarms of particle concentration (position) vector denoted as X to search for the optimum subset of features by referring to the fitness values of candidate solutions. Each sub-swarm of DEOA has identical search mechanisms with those of the single-swarm EOA. In contrary to single-swarm EOA, the proposed DEOA can leverage the advantage of an external shared pool to facilitate the sharing of equilibrium state experiences between different subswarms, enabling their particles to approach the exact Pareto front more effectively.

The search procedures of the proposed DEOA are described as follows. At the beginning of the optimization process, the initial position of each i-th particle, i.e., X_i for $i = 1, 2, \ldots, n_{\mathrm{p}}$, can be generated as follow:

$$X_{initial} = rand\left(n_p, d\right) \times (ub - lb) + lb \tag{21}$$

where n_p is the population size, d is the dimensional size of a given problem; lb and ub denote the lower and upper boundary values of decision variables, respectively. After completing the initialization process, the four best equilibrium particles (i.e., $\mathrm{X}_{eq,1}, \mathrm{X}_{eq,2}, \mathrm{X}_{eq,3}, \mathrm{X}_{eq,4}$) and the average position (i.e., $\mathrm{X}_{eq,av}$) of the population are identified to construct an equilibrium pool $\mathrm{X}_{\mathrm{eq,pool}}$ that can offer several promising search patterns as follows:

$$\mathrm{X}_{\mathrm{eq,pool}} = (\mathrm{X}_{eq,1}, \mathrm{X}_{eq,2}, \mathrm{X}_{eq,3}, \mathrm{X}_{eq,4}, \mathrm{X}_{eq,av}) \tag{22}$$

For each iteration, the original position X_{old} of each particle in every sub-swarm is updated by interacting with a solution member X_{eq} randomly selected from an equilibrium pool $X_{eq,pool}$. The solution updating mechanisms of each particle can be achieved as follows:

$$X_{new} = X_{eq} + \frac{G}{\lambda}(1-F) + (X_{old} - X_{eq}) \times F \quad (23)$$

$$F = a_1 sign\,(r-0.5)\left(e^{-\lambda t} - 1\right) \quad (24)$$

where X_{old} and X_{new} are the current and new position vectors of particles, respectively; r is the random number with a value ranges between 0 and 1; a_1 is constant (i.e., $a_1 = 2$); λ is a random vector with the value ranges between 0 and 1. An iteration counter t can be calculated as follow:

$$t = \left(1 - \frac{T}{T_{max}}\right)^{a_2(T/T_{max})} \quad (25)$$

where a_2 is constant (i.e., $a_2 = 1$); T_{max} and T denote the maximum iteration number and current iteration counter, respectively. The generation rate G can be obtained as:

$$G = \begin{cases} 0.5\, r_1 & \text{if } r_2 \geq GP \\ 0 & \text{if } r_2 < GP \end{cases} \quad (26)$$

where r_1 and r_2 denote the random numbers with range between 0 and 1; GP denotes a generation probability and it is set as 0.5.

The proposed DEOA is modified from EOA through the incorporation of external archive dominance criteria to find an appropriate set of solutions in tackling the multiobjective optimization problems that are generally described as:

$$\text{Min } F(X) = \{f_1(x), f_2(x), \ldots, f_n(x)\}$$
$$\text{Subjected to: } \begin{cases} g_i(X) \leq 0 & i = 1, 2, \ldots - q \\ h_i(X) = 0 & i = 1, 2, \ldots l \end{cases}, \quad (27)$$

where $F(x)$ is the vector of multiobjective functions; $g_i(X)$ and $h_i(X)$ denote the q inequality constraints and l equality constraints, respectively. Fig. 3 shows the flowchart of the proposed deep equilibrium optimizer algorithm. In every iteration, a new set of solutions are identified and updated in the external archive. This archive updating process enables the exchange of useful information between particles during the optimization process.

The solution quality of each equilibrium particle is evaluated iteratively by using the multiobjective functions to determine the optimal feature set from all generated features [59]. In this paper, two objective functions of minimizing the classification error ER $(X)$and minimizing the number of selected features $S(X)$ are formulated to solve the feature selection problems as follows:

$$\text{Minimize: ER}(X), S(X) X \in R^n \quad (28)$$

Without loss of generality, the fitness function used to evaluate the quality of each particle can be simplified as following [60], [61]:

$$\downarrow Fit = \alpha \text{ER} + \beta\left(\frac{|S|}{|O|}\right) \quad (29)$$

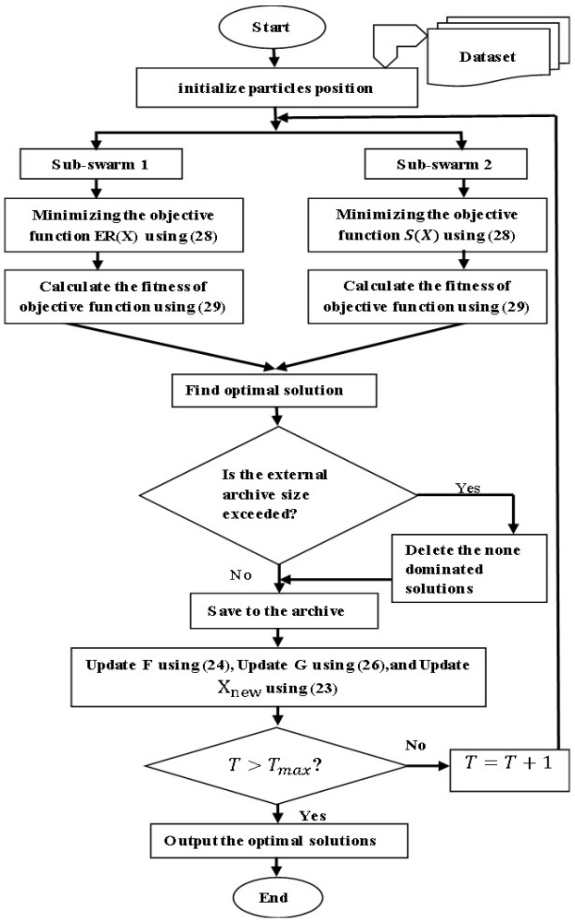

FIGURE 3. The proposed deep equilibrium optimizer algorithm (DEOA).

where $|S|$ and $|O|$ denote for the sizes of selected feature subset and original feature set, respectively; α and β are two weight infectors used to indicate the influences of classification error and size of the selected feature on the fitness function, where $\alpha \epsilon [0, 1]$ and $\beta = 1 - \alpha$ [62], [63].

D. TIME SERIES FAULT DETECTION

In this subsection, an LSTM is adopted by the proposed hybrid DL model to perform the time-series fault classification and detection of a PV system. LSTM is chosen due to the presence of its recurrent connections in the network that facilitates the memorization of data received earlier. This characteristic enables LSTM to learn the long-term time dependencies, hence can overcome the drawbacks of gradient vanishing and backflow [64]. RNNs are offered to evaluate the vague in sequential patterns of the temporal and spatial sequential data [65]. Furthermore, the connections of the peephole also allow LSTMs to identify the timed patterns accurately and compute the internal states in the cost and weight matrices [66].

Fig. 4 shows the multilayers' structure of LSTM employed to enhance the performance of the deep neural network, where the activation data from the first layer is fed to the second layer for further processing in order to deal with time

sequence problems. By connecting the LSTM layers, every layer in the LSTM is a hierarchy, which obtains input from the hidden state of the earlier layer. The sequence patterns of time series can be recognized by the training of multilayer LSTM. Therefore, the structure of connected multi-memory cells was introduced to recognize the dependency and long-term sequence of time series. As shown in Fig. 4, layer 1 of multilayer LSTM takes input from data C_{t-1}, whereas the input of layer two is obtained from its earlier time step of $h_{t-1}^{(2)}$ And the output of the present time step of layer one, i.e., $h_t^{(1)}$.

The mechanisms of LSTM cells can be explained through the mathematical formulations presented in (30) to (35). For the first layer, the forget gate of the input x_t at time t can be determined as:

$$\alpha_t = \sigma(\mathrm{W}_\alpha . [h_{t-1}, x_t] + \mathrm{b}_\alpha) \tag{30}$$

where $[\cdot]$ refers to a concatenate operation; W_α and σ denote for the weight matrix of α-th layer and the sigmoid function, respectively. Denote β_t and γ_t as the input gate layer and tanh layer, respectively, where:

$$\beta_t = \sigma(\mathrm{W}_\beta . [h_{t-1}, x_t] + \mathrm{b}_\beta) \tag{31}$$

$$\gamma_t = \tanh(\mathrm{W}_\gamma . [h_{t-1}, x_t] + \mathrm{b}_\gamma) \tag{32}$$

Meanwhile, the current state C_t can be updated from the previous state C_{t-1} as:

$$\mathrm{C}_t = \alpha_t . \mathrm{C}_{t-1} + \beta_t . \gamma_t \tag{33}$$

The output gate of the sigmoid function layer denoted as o_t, it can be computed as:

$$o_t = \sigma(\mathrm{W}_\mathrm{o} . [h_{t-1}, x_t] + \mathrm{b}_\mathrm{o}) \tag{34}$$

Referring to C_t and o_t, the current hidden layer state of h_t can be calculated as:

$$h_t = o_t \tanh(\mathrm{C}_t) \tag{35}$$

where the related layer biases are represented as b_α, b_β, b_γ, and b_o ; the related layer weights are represented as W_α, W_β, W_γ, and W_o. Finally, the final state of the LSTM network can be determined by using the $softmax(\cdot)$ activation function described as follows:

$$h_{\mathrm{final}} = softmax(h_\mathrm{t}) \tag{36}$$

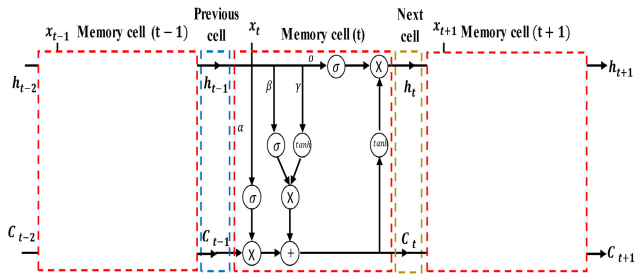

FIGURE 4. LSTM Layer for time series.

IV. CASE STUDY

A. DESCRIPTION OF PV SYSTEM

In this study, a 250 kW (kilowatt) grid-connected photovoltaic (GCPV) system adopted from [23] is considered and modeled using Matlab (2019) for performance evaluation purposes. This system consists of a PV array linked to a 25 kV utility grid with a three-phase DC/AC inverter, an inverter choke, a small harmonic filter, and a setup transformer valued at 25 kVA (Kilovolt-ampere) with the frequency of 60 Hz. As shown in Fig. 5, there is a point of common coupling (PCC) that connects the PV array to a 25 kV(kilovolt) electrical grid part. This electrical grid consists of two feeder lines 1 and 2, with lengths of 14 km(kilometer) and 8 km, respectively, connected to the PCC. The remaining part of this electrical system includes a ground transformer that is installed between Line 1 and Transformer T1. More technical details of this inverter control system operation are described in [23]. The descriptions of the considered electrical system in the current case study are presented in Table 1.

B. DATA SET GENERATIONS

A dataset of signal grid faults can be obtained from three main sources, including the operative requirements in IEEE 1547 standard, the practicable topology of the operational network, and testing practices proposed by protection relay manufacturers. All faulty cases are simulated using Matlab (R2019a). Notably, all faulty cases are generated at random times. For example, the grid faults can occur at the time period of $t = 0.3$ s, and these faults will be significantly cleared after 150 ms (millisecond). The sampling frequency and sampling time are set as 19.8 kHz and 1.5s, respectively. Table 2 summarizes the fault conditions considered in this study. Accordingly, a total of 264 fault cases are generated from datasets based on different operating conditions that include 11 types of grid faults, three types of fault locations, four types of fault resistance values, and two types of operating modes. The voltage signals were analyzed. The purpose of the analysis is to create the estimated power of the inspected PV module. The analysis of the voltage signals can be used to estimate the error among the simulated PV records compared to the threshold voltage.

V. RESULTS AND DISCUSSION

A. DATA PREPROCESSING USING WPT

In this step, WPT is applied to preprocess the PV voltage signals and analyze the incoherence of the signal in order to discover the important features used for PV fault detection. The time-frequency functions of PV signals can be obtained using (1)-(8) based on the fault conditions considered in the current study, as summarized in Table 2. Feature vectors are subsequently created from the decomposition process and fed into the network inputs of SAE in order to extract the fault features automatically.

Fig. 6 illustrates the voltage waveform observed from the case of an ABC-G fault with the resistance of 200 ohm that

TABLE 1. Descriptions of studied electrical system.

Equipment	Descriptions
Grid	120 kV, 2500MVA source
Transformer T1	120 kV/25 kV, 47 MVA(megavolt-ampere)
Transformer T2	25 kV/0.48 kV
Grounding	0.025 Ω for resistance of zero-sequence, 0.75 for reactance of zero-sequence
Distribution generation (DG)	Sun-power type (SPR-415E-WHT-D), 88 parallel strings, 7-madule. 250-kW for output power,480 V for DC source voltage, 60 Hz with 2 kHz for carrier frequency. proportional gains are 2 for voltage and 0.3for current. Integral gains are 400 for voltage and 20 for current.
Transmission line	R=3.75E-04Ω, C=0.8F, L=9.935E-05H, Line 1 is 250 kW, 25 kV with 14 km length, Line 2 is 2 MW, 25 kV with 8 km length, L3=30 MW+2 MVar.

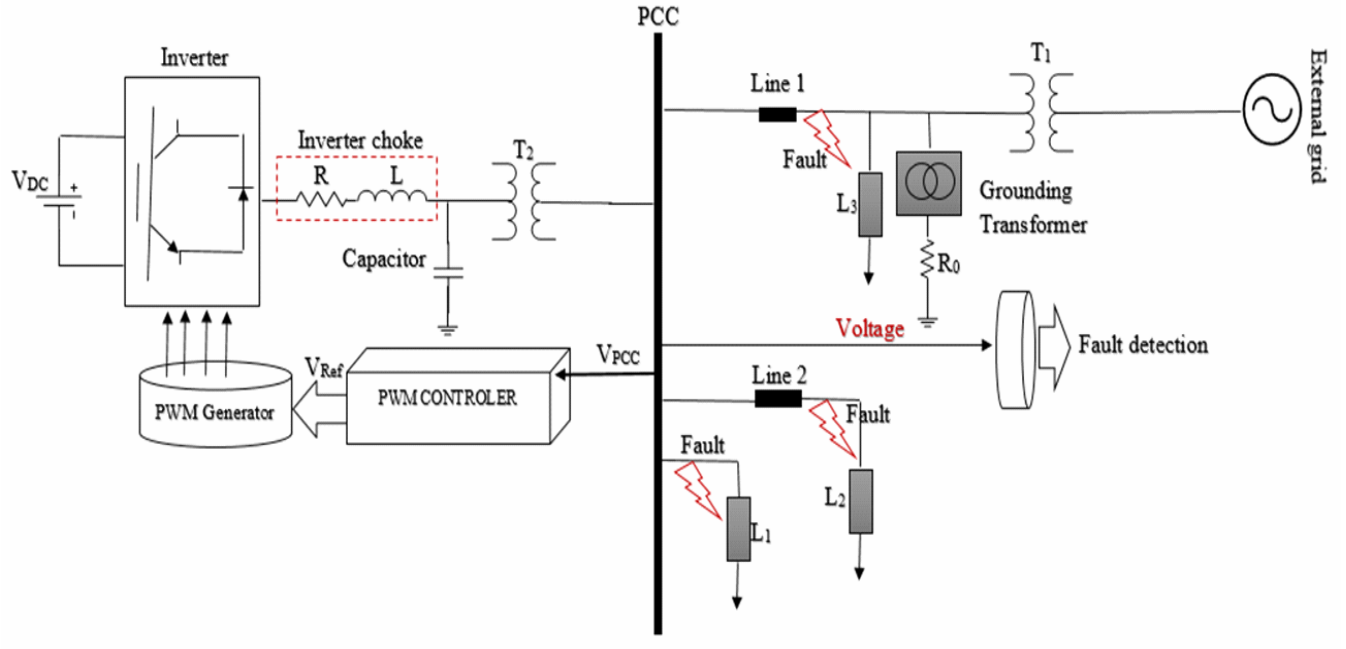

FIGURE 5. The considered electrical system in the current case study [23].

TABLE 2. Overall dataset generation.

Conditions	Descriptions	Number
Grid fault type	ABC, ABC-G, AB, BC, AC, A-G, B-G, CG, ABG, BCG, and AC-G	11
Fault location	At the PCC location, at 8 km far from the PCC, at 14 km far from the PV system.	3
Fault resistance	3, 10,75, and 200 ohms	4
Mode of operation	Islanded and grid connected	2
Total Faulty Cases		264

occurs at 14 km far from the PCC in Line 1. It is noteworthy that negligible variations of voltage amplitude are observed during the faulty event, and this undesirable scenario could be challenging for protection devices to detect the fault. Meanwhile, the effective values of voltage for different types of faulty events such as A-G, AB, AB-G, and ABC-G are also measured. The notable changes in effective voltage values can be observed at different phases for different faulty events, as shown in Fig. 7. In particular, the amplitudes of the effective voltage signal in Phase A have decreased significantly for most of the faulty events except for the AB fault. The amplitudes of effective voltage signals in Phase B are increased during the A-G fault, whereas the remaining fault events have shown a drastic drop of effective voltage signals. For Phase C, the amplitudes of effective voltage signals are increased for two cases of A-G and AB-G faults, but the decreasing of effective voltage signals is observed from the remaining faulty events. Notably, the variations of effective voltage signal can exceed the adjusted range of protection relays operation and result in the malfunction of relays. From Fig. 7, it is evident that the effective value of the voltage signal measured in the time domain can be considered as one of the fault features used to characterize the operating conditions of PV systems.

Fig. 8 illustrates the changes of frequency under different fault conditions at different locations. Accordingly, some of the faulty events, such as the ABC-G fault that occurs at

different locations, are observed to violate the upper limit of frequency constraint imposed by the grid standards. Although the presence of AB-G, AB, and A-G faults also increase the frequency value, no violation of frequency constraints is observed from these fault events. In order to prevent the malfunctions of relays during the faults, a detection method is required to quickly detect and classification of grid faults in accurate and robust manners.

B. PERFORMANCE COMPARISON ON THE EXTRACTION AND SELECTION OF FAULT FEATURES

Feature extraction is one of the most significant steps used to govern the performance of a deep learning model. In this paper, an unsupervised learning approach known as SAE is employed to extract the fault features of a given voltage signal automatically. Proper selection of network structure for SAE is influenced by various factors such as the dimension of data and the types of preprocessing techniques used. The network structure of deep SAE considered in this study consists of five hidden layers (200-100-30-10-5) with the learning rate (η = 0.3-0.3-0.3-0.3-0.2) and sparsity penalty (θ_i = 0.001, 0.154, 0.131, 0.05, 0.202) to obtain the fault features from unlabeled measured voltage signals [51]. In order to ensure good accuracy of SAE, a trial and error approach is used to determine its numbers of hidden layers [26].

When SAE is applied as the powerful feature extraction method, a large feature set consisting of 69 features is obtained. It is noteworthy that not all extracted features have high discriminatory impacts on the detection and classification performances of the proposed hybrid DL model. The presence of redundant and irrelevant features may consume longer training time but only produce the models with lower classification accuracies. Despite suffering from these technical drawbacks, most previous fault detection studies did not carefully address the feature selection issue [59]. In order to overcome these challenges, a deep equilibrium optimizer algorithm (DEOA) is proposed to select the optimal combinations of feature subsets from all extracted features.

The abovementioned feature selection problem described in (28) and (29) is solved using the proposed DEOA by minimizing the classification error and number of selected features simultaneously. The convergence curves produced by the proposed DEOA, EOA [67], and MOPSO [68] in solving the feature selection problem are illustrated in Fig. 9. Accordingly, the proposed DEOA is able to obtain better fitness values than the other two competing algorithms after 20 iterations, implying the effectiveness of the proposed algorithm in searching for better feature subsets within a shorter timeframe. Furthermore, DEOA is also observed to have a more competitive convergence speed than EOA and MOPSO because the proposed algorithm is able to converge to the best fitness function values at the iteration numbers of 85. On the other hand, the fitness values obtained by both EOA and MOPSO at the same iteration numbers are still much larger (i.e., worse) than that of DEOA. These observations imply that the optimal combinations of feature subsets obtained by DEOA have more promising quality than those of EOA and MOPSO.

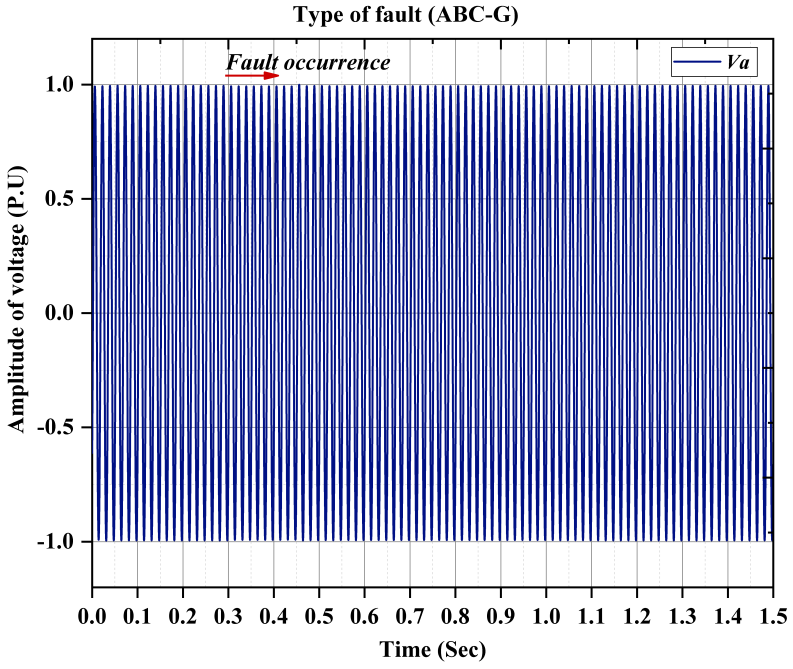

FIGURE 6. The amplitude of voltage for an ABC-G fault event that occurs at 14 km far from PCC.

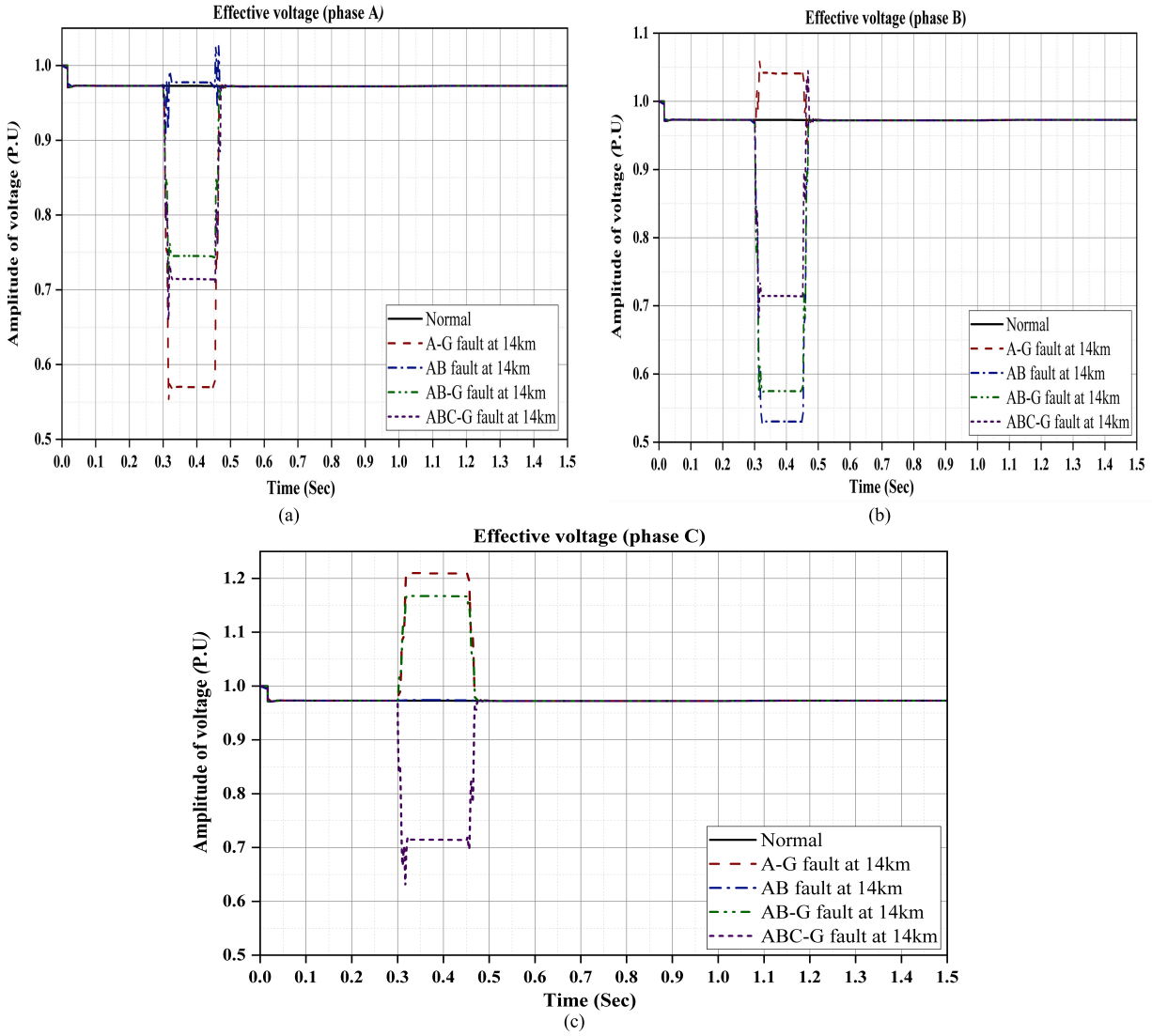

FIGURE 7. The effective values of voltage for different types of fault events in (a) Phase A, (b) Phase B, and (c) Phase C.

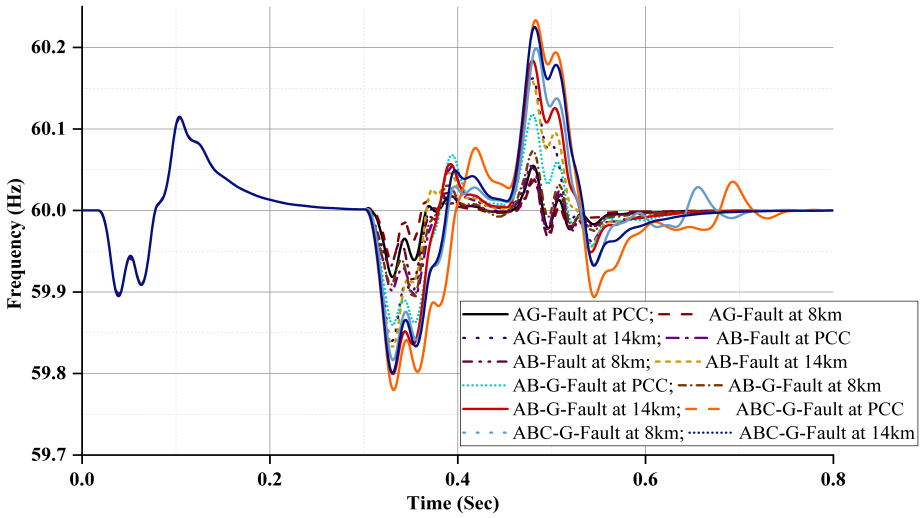

FIGURE 8. The change of frequency during fault conditions in different locations.

C. PERFORMANCE COMPARISON ON THE FAULT DETECTION AND CLASSIFICATION

In this subsection, an LSTM model is implemented together with the proposed DEOA to select the optimal feature subset required for learning the long-term date dependencies before detecting and classifying the PV faults precisely. A total of 2,000 data samples have been collected to solve the PV fault detection and classification problems, where 80% of these samples are employed as training datasets, and 20% of samples are used as the testing dataset. The training of the offered DL model is executed with TensorFlow-GPU on an NVIDIA-GeF-RTX2060 and Core i7/4.7+8700GHz (gigahertz). The variation of classification accuracy produced by the proposed hybrid DL model throughout the training and validation processes is illustrated in Fig. 10.

In order to assess the performances of the proposed hybrid DL model in detecting and classifying the PV faults, a total of 264 cases of fault events are simulated by varying the fault types, locations, resistances, and mode of operation as described in Table 2 . The overall confusion matrix produced by the proposed hybrid DL model when handling the noiseless data is presented in Table 3 . Accordingly, the proposed hybrid DL model has successfully achieved an overall accuracy of 99.93%, implying its effectiveness in detecting and classifying the PV faults for noiseless data.

The resistance of fault is one of the elements disturbing the accuracy of fault location. For the influence of the fault resistances, the simulation for different faults with fault locations and resistances was performed. The percentages of error for various faults at different resistances and locations are shown in Table 4 . The results confirm that fault distance has little effect on the proposed model even if resistance is greater than 75 Ω. The low percentages of error reported imply the high accuracy of the proposed model.

From a practical point of view, the signals measured from an actual PV system contain random noises due to the influences of electromagnetic on distribution lines. Therefore, the performance of the proposed hybrid DL model to detect and classify PV faults under noisy environments is further investigated in this subsection. Particularly, the white Gaussian noises are added to the measured voltage signals in order to simulate three different noise levels of 20 dB, 30 dB, and 40 dB when evaluating the performance of the proposed hybrid DL model to detect and classify PV faults. Table 5 shows the computational time, classification accuracy, and error values produced by the proposed hybrid DL model under different levels of noise environments. The benefits of incorporating the proposed DEOA to select the optimal subset of fault features for the proposed DL model are also further investigated in Table 5 . Accordingly, the proposed hybrid DL model has shown good robustness towards the noise environment because it is able to detect and classify the PV faults with an accuracy rate higher than 90% for different noise levels regardless of the presence of DEOA in searching for optimal feature subsets. Furthermore, the notable performance gains in terms of computing time, classification

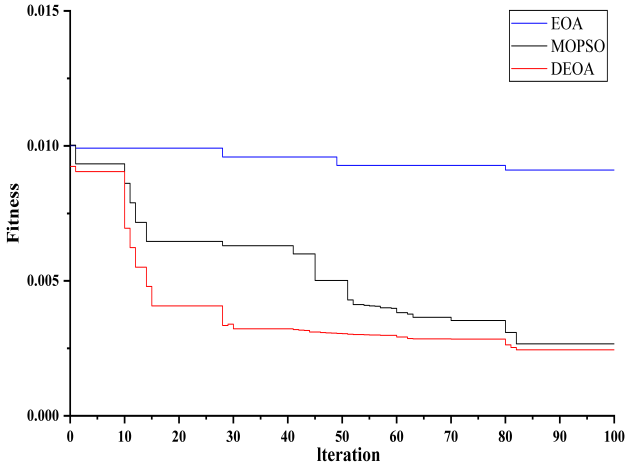

FIGURE 9. Comparison of convergence curves produced by the EOA, MOPSO, and DEOA.

accuracy, and errors are also demonstrated by the proposed hybrid DL model when DEOA is incorporated to tackle the feature selection issue. Particularly, DEOA is able to reduce at least 76% of redundant features for the proposed hybrid DL model in detecting and classifying the PV faults with a minimum accuracy level of 96.57%. The computational time incurred by the proposed hybrid DL model to perform the PV fault detection and classification is less than 0.4 seconds due to the significant reduction of fault features.

The effectiveness of the proposed hybrid DL model in detecting and classifying PV faults is further investigated through the performance comparisons with three other machine learning models, namely the LSTM [69], SAE [25], and the multilayer perceptron (MLP) [11]. The accuracy and computational time produced by all compared methods in detecting and classifying the PV faults with different numbers of selected features (i.e., all 69 features, 16 features, and 8 optimal features) are presented in Table 6 . The proposed

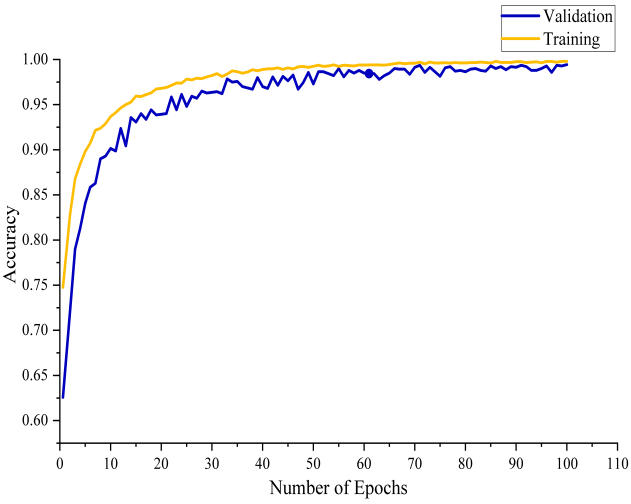

FIGURE 10. The training and validation accuracy of the proposed DL model.

TABLE 3. Overall confusion matrix of the proposed model for noiseless data.

	ABC	ABC-G	AB	BC	AC	A-G	B-G	CG	ABG	BCG	AC-G
ABC	264	0	0	0	0	0	0	0	0	0	0
ABC-G	1	263	0	0	0	0	0	0	0	0	0
AB	0	0	264	0	0	0	0	0	0	0	0
BC	0	0	0	264	0	0	0	0	0	0	0
AC	0	0	0	0	264	0	0	0	0	0	0
A-G	0	0	0	0	0	264	0	0	0	0	0
B-G	0	0	0	0	0	0	264	0	0	0	0
CG	0	0	0	0	0	0	0	264	0	0	0
ABG	0	0	0	0	0	0	0	0	264	0	0
BCG	0	0	0	0	0	0	0	0	0	264	0
AC-G	0	0	0	0	0	0	1	0	0	0	263
Overall Accuracy = 99.93%											

TABLE 4. The percentage of error for various faults at different resistances and locations.

Fault Resistance (ohms)	Fault Distance (km)	Percentage of error for different faults										
		ABC	ABC-G	AB	BC	AC	A-G	B-G	CG	ABG	BCG	AC-G
3	At the PCC location	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	At 8 km far from the PCC	0.03	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
	At 14 km far from the PV system	0.11	0.11	0.10	0.10	0.10	0.09	0.09	0.09	0.08	0.08	0.08
10	At the PCC location	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.02	0.02	0.02
	At 8 km far from the PCC	0.03	0.03	0.02	0.02	0.04	0.05	0.05	0.05	0.04	0.04	0.04
	At 14 km far from the PV system	0.35	0.35	0.20	0.20	0.20	0.14	0.14	0.14	0.20	0.20	0.20
75	At the PCC location	0.01	0.01	0.01	0.01	0.01	0.03	0.03	0.03	0.02	0.02	0.02
	At 8 km far from the PCC	0.03	0.03	0.02	0.02	0.02	0.05	0.05	0.05	0.04	0.04	0.04
	At 14 km far from the PV system	0.35	0.35	0.20	0.20	0.20	0.14	0.14	0.14	0.20	0.20	0.20
200	At the PCC location	0.05	0.05	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.04
	At 8 km far from the PCC	0.08	0.08	0.08	0.08	0.08	0.09	0.09	0.09	0.08	0.08	0.08
	At 14 km far from the PV system	0.96	0.96	0.62	0.62	0.62	0.44	0.44	0.44	0.64	0.64	0.64

TABLE 5. Experimental classification.

Noise SNR (dB)	Classification with DEOA					Classification without DEOA		
	Selected Features	Time (s)	Accuracy (%)	Error	No. of Selected Features	Time (s)	Accuracy (%)	Error
Noiseless	f2, f3, f7, f19, f24, f46, f52, f60	0.161	99.93	0.0007	8	1.347	99.53	0.0047
20	f1,f2,f3,f4,f7,f14,f17,f19,f23,f24,f31,f40,f46,f52,f60,f62	0.367	96.57	0.0343	16	1.767	91.57	0.0843
30	f1,f2,f3,f4,f7,f14,f19,f23,f24,f31,f40,f46,f52,f60	0.284	97.37	0.0263	14	1.448	95.24	0.0476
40	f2,f3,f4,f7,f14,f19,f24,f31,f46,f52,f60	0.208	99.49	0.0051	11	1.426	98.59	0.0141

TABLE 6. Comparison of results between the proposed model and other models.

Models	Training/testing data set	Classification accuracy (%)	Computational Time (s)		
			All features	16 features	Optimal features
LSTM [69]	80/20	95.31	2.126	0.4721	0.1656
SAE [25]	80/20	97.52	2.136	0.7152	0.1693
MLP [11]	80/20	92.91	6.025	2.5634	1.5454
Proposed	80/20	99.93	1.347	0.367	0.161

hybrid DL model is observed to produce the best accuracy rate of 99.93%, followed by SAE, LSTM, and MLP with the accuracy rates of 97.52%, 95.31%, and 92.91%, respectively. Although the computational times incurred by all

compared methods to detect and classify PV faults increase with the number of features considered, the proposed hybrid DL model is observed to incur the shortest computational times for all cases, followed by the LSTM, SAE, and MLP. The excellent performance of proposed hybrid DL model, in terms of accuracy rate and computational time, can be justified through the automatic feature extraction and reduction achieved by the SAE and DEO, respectively. These mechanisms enable the proposed hybrid DL model to focus on essential features and exclude those redundant ones without jeopardizing its capability to perform PV fault detection and classification. From a practical application point of view, the automatic feature extraction and reduction process is crucial to reduce the computational speed and memory consumption of processing equipment to be implemented with the proposed hybrid DL model. The excellent capabilities of the proposed hybrid DL model to perform PV fault detection and classification with high accuracy rate and short computational time under noisy environments are expected to benefit the young and experienced electrical engineers in protecting the equipment during maintenance activities.

After detecting and classifying the faults successfully using the proposed DL hybrid model, the decision-makers can select the appropriate maintenance tasks by referring to the final classification results. These classification results are expected to offer the decision-makers a clear vision of the drawbacks of a PV system and then make adequate decisions systematically. As reported by [39], CBM (condition-based maintenance) task is assigned for high-risk failure, PSM (preventive scheduled maintenance) for medium risk, and CM (corrective maintenance) for low risk.

VI. CONCLUSION AND FUTURE WORK

Manual feature extraction and feature selection are the main challenges of PV fault detection and classification because the huge feature database consists of different sensing signals under a noisy environment. In order to tackle the PV fault detection and classification problems effectively, a hybrid DL model is proposed in this paper through the proper combination of discrete wavelet transform (DWT), stacked autoencoders, deep equilibrium optimization algorithm (DEOA), and long short-term memory (LSTM). In contrary to most existing works, the proposed hybrid DL model is able to perform the automatic feature extraction and feature selection via SAE and DEOA, respectively, in order to determine the optimal feature subsets that can play decisive roles in detecting and classifying PV faults. Extensive performance analyses have demonstrated the excellent capability of the proposed hybrid DL model to solve the PV detection and classification problems with good accuracy and short computational time. Furthermore, the proposed hybrid DL model also shows its good robustness under noisy environments. The competitive fault detection and classification performances demonstrated by the proposed hybrid DL model are anticipated to benefit the electrical engineers in diagnosing the healthy conditions of PV plants during maintenance activities.

As the extended works of current studies, it is worth examining further the performance of the proposed hybrid DL model with different datasets and different multiobjective optimization problems. The same hybrid DL model can also be applied by the decision-makers in other hazardous areas such as the nuclear and gas electrical plants to evaluate and classify the risk levels in order to prevent future failures.

Motivated by the limitation of the current study, the fault location will be considered by a new method inspired by the proposed DL model and SCADA system. This method is able to detect the location of faults accurately and use to design an auto-protection control process to determine the location of simultaneous faults.

Finally, it is also possible to improve the existing hybrid DL model for online fault diagnosis in order to further reduce the undesirable effects of abnormal disturbances during the fault diagnosis.

REFERENCES

[1] S. D. Ahmed, F. S. M. Al-Ismail, M. Shafiullah, F. A. Al-Sulaiman, and I. M. El-Amin, "Grid integration challenges of wind energy: A review," *IEEE Access*, vol. 8, pp. 10857–10878, 2020, doi: 10.1109/ACCESS.2020.2964896.

[2] J. W. Burnett and F. Hefner, "Solar energy adoption: A case study of south Carolina," *Electr. J.*, vol. 34, no. 5, Jun. 2021, Art. no. 106958, doi: 10.1016/j.tej.2021.106958.

[3] M. K. H. Rabaia, M. A. Abdelkareem, E. T. Sayed, K. Elsaid, K.-J. Chae, T. Wilberforce, and A. G. Olabi, "Environmental impacts of solar energy systems: A review," *Sci. Total Environ.*, vol. 754, Feb. 2021, Art. no. 141989, doi: 10.1016/j.scitotenv.2020.141989.

[4] M. Xu, P. Xie, and B.-C. Xie, "Study of China's optimal solar photovoltaic power development path to 2050," *Resour. Policy*, vol. 65, Mar. 2020, Art. no. 101541, doi: 10.1016/j.resourpol.2019.101541.

[5] J. Li, S. Chen, Y. Wu, Q. Wang, X. Liu, L. Qi, X. Lu, and L. Gao, "How to make better use of intermittent and variable energy? A review of wind and photovoltaic power consumption in China," *Renew. Sustain. Energy Rev.*, vol. 137, Mar. 2021, Art. no. 110626, doi: 10.1016/j.rser.2020.110626.

[6] F. Aziz, A. U. Haq, S. Ahmad, Y. Mahmoud, M. Jalal, and U. Ali, "A novel convolutional neural network-based approach for fault classification in photovoltaic arrays," *IEEE Access*, vol. 8, pp. 41889–41904, 2020, doi: 10.1109/ACCESS.2020.2977116.

[7] H. Zhu, S. A. Z. Ahmed, M. A. Alfakih, M. A. Abdelbaky, A. R. Sayed, and M. A. A. Saif, "Photovoltaic failure diagnosis using sequential probabilistic neural network model," *IEEE Access*, vol. 8, pp. 220507–220522, 2020, doi: 10.1109/ACCESS.2020.3043129.

[8] M. Hajji, M.-F. Harkat, A. Kouadri, K. Abodayeh, M. Mansouri, H. Nounou, and M. Nounou, "Multivariate feature extraction based supervised machine learning for fault detection and diagnosis in photovoltaic systems," *Eur. J. Control*, vol. 59, pp. 313–321, May 2021, doi: 10.1016/j.ejcon.2020.03.004.

[9] A. F. Bendary, A. Y. Abdelaziz, M. M. Ismail, K. Mahmoud, M. Lehtonen, and M. M. F. Darwish, "Proposed ANFIS based approach for fault tracking, detection, clearing and rearrangement for photovoltaic system," *Sensors*, vol. 21, no. 7, p. 2269, Mar. 2021, doi: 10.3390/s21072269.

[10] J. Vicente-Gabriel, A. B. Gil-González, A. Luis-Reboredo, P. Chamoso, and J. M. Corchado, "LSTM networks for overcoming the challenges associated with photovoltaic module maintenance in smart cities," *Electron.*, vol. 10, no. 1, pp. 1–16, 2021, doi: 10.3390/electronics10010078.

[11] A. Ul-Haq, H. F. Sindi, S. Gul, and M. Jalal, "Modeling and fault categorization in thin-film and crystalline PV arrays through multilayer neural network algorithm," *IEEE Access*, vol. 8, pp. 102235–102255, 2020, doi: 10.1109/ACCESS.2020.2996969.

[12] A. Eskandari, J. Milimonfared, and M. Aghaei, "Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on I–V characteristics," *Sol. Energy*, vol. 211, pp. 354–365, Nov. 2020, doi: 10.1016/j.solener.2020.09.071.

[13] M. Ahmadipour, H. Hizam, M. L. Othman, and M. A. Radzi, "Islanding detection method using ridgelet probabilistic neural network in distributed generation," *Neurocomputing*, vol. 329, pp. 188–209, Feb. 2019, doi: 10.1016/j.neucom.2018.10.053.

[14] M. Ahmadipour, H. Hizam, M. L. Othman, M. A. M. Radzi, and N. Chireh, "A fast fault identification in a grid-connected photovoltaic system using wavelet multi-resolution singular spectrum entropy and support vector machine," *Energies*, vol. 12, no. 13, p. 2508, Jun. 2019, doi: 10.3390/en12132508.

[15] H. Mirshekali, R. Dashti, A. Keshavarz, A. J. Torabi, and H. R. Shaker, "A novel fault location methodology for smart distribution networks," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1277–1288, Mar. 2021, doi: 10.1109/TSG.2020.3031400.

[16] Y. D. Mamuya, Y.-D. Lee, J.-W. Shen, M. Shafiullah, and C.-C. Kuo, "Application of machine learning for fault classification and location in a radial distribution grid," *Appl. Sci.*, vol. 10, no. 14, p. 4965, Jul. 2020, doi: 10.3390/app10144965.

[17] M. Shafiullah, M. Abido, and Z. Al-Hamouz, "Wavelet-based extreme learning machine for distribution grid fault location," *IET Gener., Transmiss. Distrib.*, vol. 11, no. 11, pp. 4256–4263, Nov. 2017, doi: 10.1049/iet-gtd.2017.0656.

[18] A. Aljohani, T. Sheikhoon, A. Fataa, M. Shafiullah, and M. A. Abido, "Design and implementation of an intelligent single line to ground fault locator for distribution feeders," in *Proc. Int. Conf. Control, Autom. Diagnosis (ICCAD)*, Jul. 2019, pp. 1–6, doi: 10.1109/ICCAD46983.2019.9037950.

[19] M. Shafiullah, M. Abido, and T. Abdel-Fattah, "Distribution grids fault location employing ST based optimized machine learning approach," *Energies*, vol. 11, no. 9, p. 2328, Sep. 2018, doi: 10.3390/en11092328.

[20] E. Gord, R. Dashti, M. Najafi, and H. R. Shaker, "Real fault section estimation in electrical distribution networks based on the fault frequency component analysis," *Energies*, vol. 12, no. 6, p. 1145, Mar. 2019, doi: 10.3390/en12061145.

[21] Z. Li, J. Li, Y. Wang, and K. Wang, "A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment," *Int. J. Adv. Manuf. Technol.*, vol. 103, nos. 1–4, pp. 499–510, 2019, doi: 10.1007/s00170-019-03557-w.

[22] M. Tovar, M. Robles, and F. Rashid, "PV power prediction, using CNN-LSTM hybrid neural network model. Case of study: Temixco–Morelos, México," *Energies*, vol. 13, no. 24, p. 6512, Dec. 2020, doi: 10.3390/en13246512.

[23] M. Ahmadipour, H. Hizam, M. L. Othman, and M. A. M. Radzi, "An anti-islanding protection technique using a wavelet packet transform and a probabilistic neural network," *Energies*, vol. 11, no. 10, p. 2701, Oct. 2018, doi: 10.3390/en11102701.

[24] B. P. Kumar, G. S. Illango, M. J. B. Reddy, and N. Chilakapati, "Online fault detection and diagnosis in photovoltaic systems using wavelet packets," *IEEE J. Photovolt.*, vol. 8, no. 1, pp. 257–265, Jan. 2018, doi: 10.1109/JPHOTOV.2017.2770159.

[25] W. Gao, R.-J. Wai, and S.-Q. Chen, "Novel PV fault diagnoses via SAE and improved multi-grained cascade forest with string voltage and currents measures," *IEEE Access*, vol. 8, pp. 133144–133160, 2020, doi: 10.1109/ACCESS.2020.3010233.

[26] X. Chen, A. Ji, and G. Cheng, "A novel deep feature learning method based on the fused-stacked AEs for planetary gear fault diagnosis," *Energies*, vol. 12, no. 23, p. 4522, Nov. 2019, doi: 10.3390/en12234522.

[27] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang, and Z. Zhu, "Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery," *IEEE Access*, vol. 5, pp. 15066–15079, 2017, doi: 10.1109/ACCESS.2017.2728010.

[28] B. K. Karmakar and A. K. Pradhan, "Detection and classification of faults in solar PV array using Thevenin equivalent resistance," *IEEE J. Photovolt.*, vol. 10, no. 2, pp. 644–654, Mar. 2020, doi: 10.1109/JPHOTOV.2019.2959951.

[29] V. S. B. Kurukuru, F. Blaabjerg, M. A. Khan, and A. Haque, "A novel fault classification approach for photovoltaic systems," *Energies*, vol. 13, no. 2, p. 308, Jan. 2020, doi: 10.3390/en13020308.

[30] F. Harrou, Y. Sun, B. Taghezouit, A. Saidi, and M.-E. Hamlati, "Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches," *Renew. Energy*, vol. 116, pp. 22–37, Feb. 2018, doi: 0.1016/j.renene.2017.09.048.

[31] Z. Yi and A. H. Etemadi, "Fault detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems," *IEEE Trans. Smart Grid*, vol. 8, no. 3, pp. 1274–1283, May 2017, doi: 10.1109/TSG.2016.2587244.

[32] H. Momeni, N. Sadoogi, M. Farrokhifar, and H. F. Gharibeh, "Fault diagnosis in photovoltaic arrays using GBSSL method and proposing a fault correction system," *IEEE Trans. Ind. Informat.*, vol. 16, no. 8, pp. 5300–5308, Aug. 2020, doi: 10.1109/TII.2019.2908992.

[33] Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, and W. Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I–V characteristics," *Appl. Energy*, vol. 204, pp. 912–931, Oct. 2017, doi: 10.1016/j.apenergy.2017.05.034.

[34] A. Betti, M. Tucci, E. Crisostomi, A. Piazzi, and S. Barmada, "Fault prediction and early-detection in PV power plants based on self-organizing maps," *Sensors*, vol. 21, pp. 1–15, Jan. 2021, doi: 10.20944/preprints202101.0632.v1.

[35] M. J. Espinosa-Gavira, A. Agüera-Pérez, J. J. G. de la Rosa, J. C. Palomares-Salas, and J. M. Sierra-Fernández, "An on-line low-cost irradiance monitoring network with sub-second sampling adapted to small-scale PV systems," *Sensors*, vol. 18, no. 10, pp. 1–12, 2018, doi: 10.3390/s18103405.

[36] A. A. Arefin, "Determining islanding operation using micro grid phasor measurement unit parameters," *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 1, pp. 97–101, Sep. 2020, doi: 10.30534/ijeter/2020/1581.12020.

[37] S. K. G. Manikonda and D. N. Gaonkar, "Islanding detection method based on image classification technique using histogram of oriented gradient features," *IET Gener., Transmiss. Distrib.*, vol. 14, no. 14, pp. 2790–2799, Jul. 2020, doi: 10.1049/iet-gtd.2019.1824.

[38] A. A. Abdelsalam, A. A. Salem, E. S. Oda, and A. A. Eldesouky, "Islanding detection of microgrid incorporating inverter based DGs using long short-term memory network," *IEEE Access*, vol. 8, pp. 106471–106486, 2020, doi: 10.1109/ACCESS.2020.3000872.

[39] M. Alrifaey, T. S. Hong, A. As'arry, E. E. Supeni, and C. K. Ang, "Optimization and selection of maintenance policies in an electrical gas turbine generator based on the hybrid reliability-centered maintenance (RCM) model," *Processes*, vol. 8, no. 6, p. 670, Jun. 2020, doi: 10.3390/PR8060670.

[40] M. Alrifaey, T. S. Hong, E. Supeni, A. As'arry, and C. Ang, "Identification and prioritization of risk factors in an electrical generator based on the hybrid FMEA framework," *Energies*, vol. 12, no. 4, p. 649, Feb. 2019, doi: 10.3390/en12040649.

[41] A. Haque, K. V. S. Bharath, M. A. Khan, I. Khan, and Z. A. Jaffery, "Fault diagnosis of photovoltaic modules," *Energy Sci. Eng.*, vol. 7, no. 3, pp. 622–644, Mar. 2019, doi: 10.1002/ese3.255.

[42] K. S. Garud, S. Jayaraj, and M. Lee, "A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models," *Int. J. Energy Res.*, vol. 45, no. 1, pp. 6–35, Jan. 2021, doi: 10.1002/er.5608.

[43] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, "Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis," *IEEE Access*, vol. 7, pp. 30089–30101, 2019, doi: 10.1109/ACCESS.2019.2902949.

[44] F. Zhou, P. Hu, S. Yang, and C. Wen, "A multimodal feature fusion-based deep learning method for online fault diagnosis of rotating machinery," *Sensors*, vol. 18, no. 10, p. 3521, Oct. 2018, doi: 10.3390/s18103521.

[45] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *IEEE Access*, vol. 6, pp. 26241–26253, 2018, doi: 10.1109/ACCESS.2018.2837621.

[46] J. Nalic and A. Svraka, "Importance of data pre-processing in credit scoring models based on data mining approaches," in *Proc. 41st Int. Conv. Inf. Commun. Technol., Electron. Microelectron. (MIPRO)*, May 2018, pp. 1046–1051, doi: 10.23919/MIPRO.2018.8400191.

[47] R. X. Gao and R. Yan, "Wavelet packet transform," in *Wavelets: Theory and Applications for Manufacturing*, R. X. Gao and R. Yan, Eds. Boston, MA, USA: Springer, 2011, pp. 69–81.

[48] S. K. Mishra and L. N. Tripathy, "A critical fault detection analysis & fault time in a UPFC transmission line," *Protection Control Mod. Power Syst.*, vol. 4, no. 1, pp. 1–10, Dec. 2019, doi: 10.1186/s41601-019-0117-5.

[49] Q. He, "Vibration signal classification by wavelet packet energy flow manifold learning," *J. Sound Vib.*, vol. 332, no. 7, pp. 1881–1894, Apr. 2013, doi: 10.1016/j.jsv.2012.11.006.

[50] C. Lu, Z.-Y. Wang, W.-L. Qin, and J. Ma, "Fault diagnosis of rotary machinery components using a stacked denoising autoencoder-based health state identification," *Signal Process.*, vol. 130, pp. 377–388, Jan. 2017, doi: 10.1016/j.sigpro.2016.07.028.
[51] M. Alrifaey, W. H. Lim, and C. K. Ang, "A novel deep learning framework based RNN-SAE for fault detection of electrical gas generator," *IEEE Access*, vol. 9, pp. 21433–21442, 2021, doi: 10.1109/ACCESS.2021.3055427.
[52] A. Faramarzi, M. Heidarinejad, B. Stephens, and S. Mirjalili, "Equilibrium optimizer: A novel optimization algorithm," *Knowl.-Based Syst.*, vol. 191, Mar. 2020, Art. no. 105190, doi: 10.1016/j.knosys.2019.105190.
[53] A. M. Shaheen, A. M. Elsayed, R. A. El-Sehiemy, and A. Y. Abdelaziz, "Equilibrium optimization algorithm for network reconfiguration and distributed generation allocation in power systems," *Appl. Soft Comput.*, vol. 98, Jan. 2021, Art. no. 106867, doi: 10.1016/j.asoc.2020.106867.
[54] S. K. Dinkar, K. Deep, S. Mirjalili, and S. Thapliyal, "Opposition-based Laplacian equilibrium optimizer with application in image segmentation using multilevel thresholding," *Expert Syst. Appl.*, vol. 174, Jul. 2021, Art. no. 114766, doi: 10.1016/j.eswa.2021.114766.
[55] S. I. Seleem, H. M. Hasanien, and A. A. El-Fergany, "Equilibrium optimizer for parameter extraction of a fuel cell dynamic model," *Renew. Energy*, vol. 169, pp. 117–128, May 2021, doi: 10.1016/j.renene.2020.12.131.
[56] J. Wang, B. Yang, D. Li, C. Zeng, Y. Chen, Z. Guo, X. Zhang, T. Tan, H. Shu, and T. Yu, "Photovoltaic cell parameter estimation based on improved equilibrium optimizer algorithm," *Energy Convers. Manage.*, vol. 236, May 2021, Art. no. 114051, doi: 10.1016/j.enconman.2021.114051.
[57] M. A. Soliman, A. Al-Durra, and H. M. Hasanien, "Electrical parameters identification of three-diode photovoltaic model based on equilibrium optimizer algorithm," *IEEE Access*, vol. 9, pp. 41891–41901, 2021, doi: 10.1109/ACCESS.2021.3065386.
[58] M. Micev, M. Ćalasan, and D. Oliva, "Design and robustness analysis of an automatic voltage regulator system controller by using equilibrium optimizer algorithm," *Comput. Electr. Eng.*, vol. 89, Jan. 2021, Art. no. 106930, doi: 10.1016/j.compeleceng.2020.106930.
[59] G. I. Sayed, A. Tharwat, and A. E. Hassanien, "Chaotic Dragonfly algorithm: An improved metaheuristic algorithm for feature selection," *Appl. Intell.*, vol. 49, no. 1, pp. 188–205, 2019, doi: 10.1007/s10489-018-1261-8.
[60] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary ant lion approaches for feature selection," *Neurocomputing*, vol. 213, pp. 54–65, Nov. 2016, doi: 10.1016/j.neucom.2016.03.101.
[61] M. M. Mafarja, D. Eleyan, I. Jaber, A. Hammouri, and S. Mirjalili, "Binary dragonfly algorithm for feature selection," in *Proc. Int. Conf. New Trends Comput. Sci. (ICTCS)*, Oct. 2017, pp. 12–17, doi: 10.1109/ICTCS.2017.43.
[62] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371–381, Jan. 2016, doi: 10.1016/j.neucom.2015.06.083.
[63] M. Mafarja, I. Aljarah, H. Faris, A. I. Hammouri, A. M. Al-Zoubi, and S. Mirjalili, "Binary grasshopper optimisation algorithm approaches for feature selection problems," *Expert Syst. Appl.*, vol. 117, pp. 267–286, Mar. 2019, doi: 10.1016/j.eswa.2018.09.015.
[64] M. A. Ranzato, C. Poultney, S. Chopra, and Y. L. Cun, "Efficient learning of sparse representations with an energy-based model," in *Proc. Adv. Neural Inf. Process. Syst.*, 2007, pp. 1137–1144.
[65] K.-I. Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural Netw.*, vol. 6, no. 6, pp. 801–806, 1993, doi: 10.1016/S0893-6080(05)80125-X.
[66] F. D. D. S. Lima, G. M. R. Amaral, L. G. D. M. Leite, J. P. P. Gomes, and J. D. C. Machado, "Predicting failures in hard drives with LSTM networks," in *Proc. Brazilian Conf. Intell. Syst. (BRACIS)*, Oct. 2017, pp. 222–227, doi: 10.1109/BRACIS.2017.72.
[67] D. Piasson, A. A. P. Bíscaro, F. B. Leão, and J. R. S. Mantovani, "A new approach for reliability-centered maintenance programs in electric power distribution systems based on a multiobjective genetic algorithm," *Electr. Power Syst. Res.*, vol. 137, pp. 41–50, Aug. 2016, doi: 10.1016/j.epsr.2016.03.040.
[68] J.-S. Chou and T.-S. Le, "Reliability-based performance simulation for optimized pavement maintenance," *Rel. Eng. Syst. Saf.*, vol. 96, no. 10, pp. 1402–1410, Oct. 2011, doi: 10.1016/j.ress.2011.05.005.
[69] V. Veerasamy, N. I. A. Wahab, M. L. Othman, S. Padmanaban, K. Sekar, R. Ramachandran, H. Hizam, A. Vinayagam, and M. Z. Islam, "LSTM recurrent neural network classifier for high impedance fault detection in solar PV integrated power system," *IEEE Access*, vol. 9, pp. 32672–32687, 2021, doi: 10.1109/ACCESS.2021.3060800.

MOATH ALRIFAEY (Graduate Student Member, IEEE) received the B.S. degree in electronic engineering from Ibb University, Yemen, in 2005, the M.S. degree in engineering management from the Delft University of Technology and Taiz University, in 2014, and the Ph.D. degree in industrial engineering from Universiti Putra Malaysia, Malaysia, in 2020.

From 2008 to 2016, he worked as a Control System and Instrumentation Engineer at the Oil and Gas Plant, Yemen. He is currently a Postdoctoral Research Fellow with the Faculty of Engineering Technology and Built Environment, UCSI University. His current research interests include intelligent manufacturing, reliable maintenance, and computational intelligence, focusing on problems of classification, optimization, and machine learning, deeply in industrial applications and big data engineering.

WEI HONG LIM received the B.Eng. degree (Hons.) in mechatronic engineering and the Ph.D. degree in computational intelligence from Universiti Sains Malaysia, Penang, Malaysia, in 2011 and 2014, respectively. He is currently an Assistant Professor and a Researcher with the Faculty of Engineering, Technology and Built Environment, UCSI University. He was attached to the Intelligent Control Laboratory, National Taipei University of Technology, Taiwan, as the Postdoctoral Researcher, from 2015 to 2017, and as the Visiting Researcher, in 2019. He has published more than 50 research articles in research areas related to computational intelligence, optimization algorithms, energy management, and digital image processing. He is currently working with two national research grants awarded by the Ministry of Education, Malaysia, and five internal grant projects supported by UCSI University. He is an Active Reviewer for various reputable journals, such as IEEE Access, *Complexity*, *Mathematical Problems in Engineering*, and *Computational Intelligence and Neuroscience*. He is also involved in various professional bodies. To date, he has been awarded with Chartered Engineer (C.Eng.) qualification from U.K. Engineering Council and Professional Technologist (P.Tech.) qualification from the Malaysia Board of Technologist (MBOT).

CHUN KIT ANG received the B.Eng. degree (Hons.) in mechatronic engineering from UCSI University, Malaysia, in 2010, and the Ph.D. degree in mechanical and manufacturing engineering from the Universiti Putra Malaysia, in 2014. He is currently the Dean of the Faculty of Engineering Technology and Built Environment, UCSI University. He has published many papers related to the application of artificial intelligence in recent years. His research interests include artificial intelligence, soft computing, robotics, and mechatronics.

ELANGO NATARAJAN is currently an Associate Professor with the Department of Mechanical and Mechatronic Engineering, UCSI University, Kuala Lumpur, Malaysia. His research interests include the machining of composite materials and optimization, machine learning in manufacturing, soft robotics, and composite materials. He has been in higher learning education for more than 20 years. He has completed many funded projects and has published 66 research articles in refereed Scopus/WoS journals as of today and his H-index is 12. He is associated with many professional bodies, including IEEE, IET, and Board of Engineers Malaysia. He has been appointed as a Professional Review Interviewer by IET, U.K. He is currently the Secretary of IEEE/Robotics and Automation Society, Malaysia. He has obtained Chartered Engineer (C.Eng.) from Engineering Council, U.K.

MAHMUD IWAN SOLIHIN is currently an Associate Professor with the Department of Mechatronics Engineering specialized in applications of artificial intelligence/machine learning, data-driven spectroscopy, meta-heuristics optimization, control systems, and robotics. In the past, he was a Research Assistant for an eScience-Fund project under the Ministry of Science Technology & Innovation (MOSTI), Malaysia. His current research interest is mainly on applications of machine learning and data-drive modeling in various fields, including agriculture, water resource management, the IoT, control systems, robotics, and sensor networks, toward sustainable engineering development in the future. He is also a Professional Member of the Institution of Engineers of Indonesia and recognized by the ASEAN Federation of Engineering Organisations (AFEO) and APEC Engineer.

MOHD RIZON MOHAMED JUHARI received the B.Eng. and M.Eng. degrees in electrical and electronics engineering from Tokushima University, Japan, in 1993 and 1995, respectively, and the D.Eng. degree from the Department of Computer Science and Intelligent System, Oita University, Japan, in 2002. From 1995 to 1997, he was a Software Engineer at the System LSI Laboratory, Mitsubishi Electric Corporation, Itami, Japan. In 1997, he moved to the University of Malaya, Malaysia. He was the Head of the Biomedical Engineering Department, University of Malaya, and was a Professor of biomedical technology at King Saud University, Saudi Arabia, in 2010. He served at MIEC, Fuzhou University, China, in 2019. He is currently a Professor with the Department of Electrical and Electronics, UCSI University, Malaysia. He has written more than 150 published journal articles and proceedings in the human face analysis area. His research interests include pattern recognition, biometrics system, face analysis, face recognition, object recognition, and vision for a mobile robot. His research invention has been utilized and modeled out into a product of face recognition system which has been coined into ''Iris-ID.'' The product value was recognized and commended with a Silver Medal Award at the international level in the 34e Salon International des Inventions des Techniques et Produits Nouveaux 2006, Geneva. He is a Professional Engineer with Practicing Certificate (Ir.), Board of Engineers Malaysia, Professional Technologist (Ts.), Malaysia Board of Technologists, Chartered Engineer (C.Eng.), Engineering Council, U.K., a Senior Member of The Institution of Engineers, Malaysia (SMIEM), and a member of the Institute of Engineering Technology (MIET), U.K.

SEW SUN TIANG received the B.Eng. degree in electronic engineering (telecommunications) from Multimedia University, Melacca, Malaysia, in 2008, and the Ph.D. degree in electrical and electronic engineering from the Universiti Sains Malaysia, Penang, Malaysia, in 2014.

From 2015 to 2017, she was a Senior Lecturer with the School of Engineering, Asia Pacific University, Kuala Lumpur, Malaysia. She is currently an Assistant Professor with the Faculty of Engineering, Technology and Built Environment, UCSI University. Her research interests include medical applications of microwaves, antenna design, and optimization.